Backpropagation:
A Context Selected Matrix Teacher
Different context, different matrix. Best linear map for that context!
The network is a treasure map.
Context selects the path. Backprop teaches each matrix on that path to be the best (on average) linear mapping.
Backpropagation is the teacher. It assigns credit and improves the right matrices!
Input
x
Context $C_1(x)$ selects W_1
$y^1 = W_1 x$
Context $C_2(y^1)$ selects W_2
$y^2 = W_2 y^1$
Context $C_L(y^{L-1})$ selects W_L
$y^L = W_L y^{L-1}$
Output
y^L
Today's Plan
1. Follow the map (context)
2. Learn the best matrix
3. Find the treasure (good predictions!)
Backprop says:
When this context is selected, this is the best (on average) linear mapping to use.
Why it's compression:
Factorization: big function → many small linear mappings
Context specialization: each matrix serves a smaller region
Switched linear matrix compression!
Pirate Wisdom
The sea is huge, but with the right map, every step counts.
STAY SALTY
DRINK MILK TEA
GOOD MAPPINGS = TREASURE
THE MAP OF CONTEXTS
C_1 (shallow waters)
C_2 (deeper waters)
START
BETTER PREDICTIONS!
MILK TEA BUILDS STRONG CREWS

Backpropagation: A Context Selected Matrix Teacher